DOCUMENT RESUME

ED 464 114 TM 033 795

AUTHOR Lambating, Julita; Allen, James D.
TITLE How the Multiple Functions of Grades Influence Their Validity and Value as Measures of Academic Achievement.
PUB DATE 2002-04-00
NOTE 32p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 1-5, 2002).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150) -- Tests/Questionnaires (160)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Attitudes; Case Studies; *College Students; *Grades (Scholastic); Grading; *High School Students; High Schools; Higher Education; *Parents; Preservice Teacher Education; Preservice Teachers; *Teachers; Validity

ABSTRACT

This study explored the perspectives of high school students, high school teachers, undergraduate and graduate students in preservice teacher education programs, and university professors about how a high school teacher assigns grades. The study also asked whether these different groups gave more or less weight to completing an assigned task than demonstrating mastery of the subject in the assignment of grades and whether a stand-alone measurement course made a difference among educators' perspectives on grading. A survey was developed to collect perspectives through a case study of a realistic grading dilemma. Data were collected from 153 preservice teachers (mostly undergraduates), 49 preservice teachers in a graduate course, 81 practicing high school and elementary school teachers in 2 states, 34 education professors, and 288 high school students in 2 states. Participants responded in a similar way to the case study, but there was a difference in the perceptions of high school students and those of other groups. Educators and students agreed that homework as for practice, but most respondents would consider homework in grading the student. Generally respondents had similar views of the purpose of a grade. Educators who had taken a measurement course were more likely to consider homework as a requirement that should affect grading. Implications for teacher education are discussed. Appended are the survey instruments. (Contains 13 tables and 40 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 464 114

# How the multiple functions of grades influence their validity and value as measures of academic achievement

Julita Lambating, Ph.D.
The California State University
Sacramento, California

James D. Allen, Ph.D.
The College of Saint Rose
Albany, New York

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J. Allen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM033795

Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans, April 2002
Session 28.71

Introduction

Grades are one of the most important educational commodities that a student obtains. However, determining the quality and worth of grades is not as simple as it seems. In fact, it is such a complex issue that some educators have proposed their abolition (Manzano, 2000; Popham, 2000). Grades can open up or close down important learning opportunities for students (Jasmine, 1999). With high grades, students get admitted to colleges and universities of their choice and receive scholarships and tuition assistance since grades are a major selection criteria in the college admission process. The reverse is also true. It is very difficult for students to get admitted to some schools if their grades are not sufficiently high. In addition, if a high grade is obtained, but not deserved, a student could be misplaced in an academic course or program for which she or he is not prepared and experience failure. Likewise, based on principles of attribution and social cognitive theories, if students receive grades lower than ones that accurately depicts their true level of academic knowledge, it may lead students to believe they lack the ability to succeed academically and lower their sense of self-efficacy as well as their motivation to learn (Pintrich & Schunk, 2002). Although grades are defined to be marks "showing the quality" of a student's academic work (American Heritage Dictionary, 1983, p. 303), that is, to what extent one has learned, grades often fail to communicate that very quality.

Confusing Purposes and Miscommunication

It is very important then that there is congruence in the understanding of the appropriate functions of grades and the information they communicate between the people who issue the grades and the ones who will interpret the grades. Absence of this coherence will lead to inappropriate use of grades minimizing their value and usefulness as basic currency in our educational system. This lack of coherence seems to be found throughout our education system.

In a study by Baron (2000), results show that there is lack of coherence in the beliefs about grades held by parents and students and the beliefs held by the education community. Even within the same school, teachers often hold very different views about the purpose of grades and fail to communicate with their colleagues about their grading practices (Kain, 1996). New teachers often work independently and are left to figure out their own grading policies, gradually adhering to the norms of the school. There is a similar lack of coherence and communication among college teachers (Barnes, Bull, Campbell & Perry, 1998).

This paper takes a strong position regarding the purpose of grades that become part of a student's permanent record such as those reported on transcripts and report cards. The purpose of these grades is to communicate a valid and reliable summary of a student's academic achievement in the subject that is listed next to the grade on the record. It can not, and therefore should not, be a hodgepodge of factors such as student's level of effort, innate aptitude, compliance to rules, social behaviors, attitudes, or other non-achievement measures. It can not be a teacher's "merged judgement[1]" of these factors, since as a single letter or numeric mark, the reported grade must communicate a single fact about the student if it is to be a valid or accurate source of information coherently shared between the reporter of the grade and the person reading the grade report. The purpose of an academic report is to communicate the level of academic achievement that a student has developed over a course of study. Therefore, the sole purpose of a grade on an academic report, if it is to be a valid source of information, is to communicate the academic achievement of the student. If other factors about the student are deemed important, such as a student's attitude, level of effort, or social behavior, then other appropriate forms of reporting these factors must be made available and used. If a multidimensional view of the student is desired, then a multidimensional system of reporting is required. Using a single grade

as a summary of a teacher's "merged judgement" of a student leads to miscommunication, confusion, and a continuation of the lack of coherence among stakeholders about what a grade represents. Due to the difficult task of aligning learning standards, instruction, and assessment, it is hard enough for teachers to ensure the validity of a grade even when related to the single factor of achievement. Including other factors can only complicate the teacher's task of communicating to others the academic progress of their students.

Stiggins (2001) provides a list of questions that a reflective teacher might ask themselves about their grading practices. Several of his questions get to the heart of the validity issue of a teacher's grading practices. These questions include:

- Why do I have to assign grades? What's the purpose?
- Should I grade just on ability and achievement or should I consider effort and attitude as well?
- Are my grades supposed to mean the same thing as those of others who teach the same courses? How can we ensure that this is true?
- If we all reflect different expectations and standards in our grades, how can anyone interpret our grades accurately? (p. 410)

Stiggins (2001) effectively argues that report card grades must be accurate communications of students' achievement and not for other purposes such as to motivate students or control their behavior, which can subvert, or invalidate, the communication. He addresses the five factors (achievement, aptitude, effort, compliance, attitudes) that teachers often merge together to calculate grades. Based on principles of measurement, learning, and motivation, he systematically eliminates all except achievement as appropriate and valid factors to include in the reporting of grades.

---

[1] The authors have borrowed this phrase from an anonymous reviewer.

Stiggins (2001) provides the following example which is especially relevant to the focus of the research presented in this paper.

> For instance, let's say a teacher uses homework grades as a motivator to compel students to practice because this teacher believes that those who do so learn more. He warns students that failure to turn in any homework assignment will be entered as a failure in the gradebook. All incomplete assignments turn into Fs when it's time to compute that final grade.
>
> Now let's say there are two students in this class, both of whom have developed a very high level of mastery of the material. Each has an outstanding record on all tests, quizzes, and projects. Yet one student has consistently failed to complete practice assignments and so has accumulated many Fs for homework in the gradebook.
>
> At report card grading time, the message sender (the teacher) sends out word (via grades) that one student has been assigned an A, while the other is assigned a C. The resulting communication problem becomes apparent when we realize that both students learned the same amount. The message receivers (parents, other teachers, etc.) have no basis on which to distinguish the messages. They won't know or understand the subtleties of meaning hidden within the message and may draw inappropriate conclusions about each student's achievement. The result is miscommunication. (p. 411)

Although report card grades should be assigned to communicate to others the level of academic achievement that a student has attained (Snowman, Biehler & Bonk, 2000), this principle is often ignored in practice (Tombari & Borich, 1999). If grades are to communicate students' academic achievement, then classroom assessments on which these grades are based should be measures of what level of subject content students have learned (Gronlund, 1998;

Oosterhof, 2001). The main goal for classroom assessment is to gather reliable, valid, meaningful, and useful information about student learning (Brookhart, 1998; Linn and Gronlund, 1995). A grade that a student gets at the end of a semester is a single numerical or letter index summarizing the results of classroom assessment. It is a summary evaluation that is used to make immediate and important decisions such as skipping some courses or taking a remedial course as well as make long range career plans (Gage and Berliner, 1992). Assigning, and therefore interpreting, this grade can be difficult because of the complexity of the teaching and learning process that is linked to assessment. It has been suggested by Brookhart (1998) that classroom assessment and grading practices are at the center of effective management of classroom instruction and learning.

Grading and Lack of Professional Training

To add to the lack of cohesion and miscommunication is the problem with pre-service teacher training at the college level (Goodwin, 2001). Pre-service teachers have been students for 16+ years before finishing their career preparation as full-time teachers. They form many of their perspectives about teaching from these many years of observing teachers and their teaching practices (Lortie, 1975). They have been recipients of hundreds, if not thousands, of grades from their K-12 teachers and college professors before taking on the responsibility of assigning grades to their own students. To a large degree their perception regarding grades come from their own long experience as students. The only formal training on grading procedures they receive before they are responsible for grading their own students comes from their teacher education programs. Very few teacher education programs (less than 1/3) include a stand-alone "Measurement" or "Assessment" course (Allen & Lambating, 2001, see Appendix A). More often than not, the topic of grades is just one of many that are addressed in a methods or instructional design course.

According to Stiggins (1999), how the concepts of "reliability" and "validity" are related to classroom practice is not addressed in the courses in which our pre-service teachers are introduced to these terms. It is important to look at this issue because the principles of validity and reliability are considered the most fundamental principles related to measurement and therefore important to classroom assessment (Gallagher, 1998; Gredler, 1999; Linn & Gronlund, 1995).

Grading systems used by teachers vary widely and unpredictably and often have low levels of validity and reliability due to the inclusion of non-academic criteria used in the calculation of grades (Allen & Lambating, 2001; Brookhart, 1994; Friedman & Frisbie, 1995; Johnsen, 1995; Olson, 1989). Grading practices by teachers rarely follow the measurement principles and the grading practices recommended in measurement textbooks (Cross and Frary, 1996). Some believe that this is due to the lack of adequate training of teachers in objective and systematic methods of assessment and evaluation leaving teachers to rely on poor grading practices (Stiggins, 1988). Others argue that even when teachers are provided with measurement instruction, they still use subjective value judgements when assigning grades (Brookhart, 1993). Research suggests that pre-service teachers need hands-on experience in ways of grading students and how to work with cooperating teachers who assess and grade in ways different from those learned from their instructors on assessment and grading (Barnes, 1985; Lomax, 1996). Several researchers have shown that grades are used for multiple and contradictory purposes, and there use as a tool for communicating student achievement is just one of these many purposes. It has been shown that grades are used as a motivational tool as well as to develop desirable behaviors such as good study habits (Oosterhof, 2001). According to Gronlund (1998), grading based on learning ability has sometimes been used at the elementary level to motivate students

with less ability. It follows then that instead of the grade being a function of what a student has learned it has become a function of many variables. Included in these variables are submissions of homework, class attendance, and non-academic behaviors. Simply put, it would appear that grades are measures of how well a student lives up to the teacher's expectation of what a good student is rather than measures of academic achievement by the student.

Considering the complexity of the issue of grades and grading procedures, this paper explores three issues:

1. How are the perspectives of high school students, full-time high school teachers, undergraduate and graduate students in pre-service teacher education programs, and education professors similar or different regarding how a high school teacher assigns grades?
2. Do high school students, full-time high school teachers, undergraduate and graduate students in pre-service teacher education programs, and education professors give more or less weight to completing an assigned task (e.g., homework) than demonstrating mastery of the subject in the assignment of grades?
3. Does taking a stand-alone measurement course make a difference among educators' perspectives on grading?

Methods of Inquiry

The use of case studies has become very common in teacher education courses including educational psychology classes (Allen, 1999; Sudzina, 1999). This is illustrated by the adoption of short cases or vignettes in almost every currently published Educational Psychology textbook. Case study analysis is particularly useful in developing problem solving skills and reflective thinking among students (Allen, 1995; Richert, 1991; Silverman, Welty & Lyon, 1996). The use of case with pre-service teachers help them to move beyond their preconceived and often

simplistic notions about teaching practices that are based on their personal experiences and to consider the complexity of teaching illustrated in well-crafted case studies linking educational theory to practice (Goodwin, 2001). In addition, the use of surveys is a commonly accepted method for collecting descriptive data in educational research (Gall, Gall & Borg, 1999; McMillan & Schumacher, 2001; Mertens, 1998). The use of a case study in conjunction with a survey instrument provided a rich context for subjects to respond to research questions in a more realistic and informative manner.

A survey was developed to collect data on educators' perspectives about assessment and grading issues as addressed in a case study (see Appendix B). The case study presents a realistic grading dilemma in which a teacher (Sarah) assigns grades based on criteria that could be considered both invalid and unreliable. The use of this case study was used for several reasons. First, over the past decade one of the authors of this research has used this case to help students in his educational psychology course reflect on the complexity of classroom assessment. Each semester it is the case that students are the most reluctant in accepting alternative perspectives on grading than the one that has been instilled in them over their careers as students, specifically the perspectives that effort and compliance to requirements should be included in determining a student's grade. Also, the research by Brookhart (1993) on teachers' grading practices involved collection of data through the use of a series of one-paragraph scenarios about teacher's grading practices. Although we found this method good at attempting to get teachers to reflect on and respond to "realistic" classroom situations, we thought that they lacked the true complexity of classrooms and the multiplicity of factors that teachers must take into consideration in making decisions. The case study of "Sarah Hanover" is four pages long and allows for that complexity to be developed (Silverman, Welty, & Lyon, 1996).

Research participants were asked to read and analyze the case and provide a solution to the grading dilemma along with rationale for their decisions. The major dilemma is that the teacher, Sarah, has given the most knowledgeable student (James) in her math class [98% test average, best peer-tutor, active participant in class, pleasant personality] a letter grade of "B" because he never turned in homework. The parents of the student want Sarah to change the student's grade to an "A" to reflect his knowledge of math. The major dilemma for Sarah is whether to change the grade from B to a higher grade as requested by James' parents. Questions of the validity of the grade and the reliability of homework are central issues in the case. Research participants were asked to respond to the following four questions about grading and homework after reading the case:

- Q1: Should Sarah change James' grade? No Yes Why/why not?
- Q2: What letter grade does James deserve? A B C D F Why?
- Q3: What do <u>you</u> believe should be the purpose of homework?
- Q4: What do <u>you</u> believe should be the purpose of a grade?

Responses to the survey questions based on the grading scenario described in the case provided the researchers greater insight on the perspectives of each of the respondents regarding grades.

<u>Subjects</u>

Data were collected from the following groups of educators and students:

- 153 pre-service teachers- the majority (151) of whom were enrolled in an undergraduate educational psychology survey course. These students were in their third year of the various teacher-education programs (El. Ed., Special Ed., Secondary Ed., Art Ed., etc.) in a School of Education at a private liberal arts college in upstate New York. These undergraduate students do not take a course in measurement. Their primary exposure to measurement theory and

grading is as one of the major topics covered in the undergraduate educational psychology survey course (approximately 2 weeks out of a 15-week semester). They had read a chapter on classroom assessment and had 2-3 class periods of lecture/discussion on the chapter before reading the case and completing the questionnaire.

- 49 pre-service teachers enrolled in either a graduate educational psychology survey course in New York (n = 37) or graduate education courses at a California State university (n = 12). The majority (n = 35) of the New York graduate students were students not having an undergraduate education degree. They were taking the survey course as a foundation for their other Master's level education courses and were being introduced to measurement theory and classroom grading issues for the first time (again in 2 weeks of a 15-week semester). The New York graduate students had read a chapter on classroom assessment and had 2-3 class periods of lecture/discussion on the chapter before reading the case and completing the questionnaire. The California graduate students were doing their student teaching experience that is part of the credential program. These students had taken an Educational Psychology course that includes approximately six hours of discussion on classroom assessment.
- 81 practicing high school and elementary teachers teaching in New York (n = 44) or California (n = 37). The majority (31) of the New York teachers indicated that they had taken a "graduate level course in tests and assessments." Seventeen of the 37 California graduate students were taking an assessment course when the data were collected.
- 34 school of education professors teaching in New York (n = 23) or California (n = 11). These included professors in departments of teacher education, special education, educational psychology, reading, communication disorders, and counseling.

- 288 high schools students in New York (n = 163) or California (n = 125). The students in New York were 10th, 11th, and 12th graders taking either an Honors or Advance Placement English or science course. The California students were 9th and 10th graders in mixed ability science classes.

The college and high school students were asked to complete the survey as a voluntary non-graded activity. In New York, high school and college teachers were provided a $5 incentive (gift certificate/movie pass) to voluntarily complete the survey. The response rate for these high school teachers was 63% (44 of 70 surveys sent) and for college faculty 51% (23 of 45 surveys).

Data Analysis

A sample of data for each of the four survey questions was read to identify categories to classify subjects' essay responses. This analysis resulted in several categories for each question included in the survey instrument (see Appendix C). Two graduate students were trained by the second author of this study to identify these major categories to classify all the responses on all surveys. If a response did not fit into any of the pre-determined categories, a new category was established if the response was relevant to the case. In the few cases that the written response was not logically related to the case study, it was not included in the data set.

The large number of individuals responding to each item in the survey as well as the variability of the groups responding to the same set of items helps to ensure reliability of the data collected. The range of variation in the categories resulting from the explanations and reasons provided by the respondents to the primary and follow-up questions adds to the reliability of the data obtained. The case study and survey was piloted with several classes of students to insure that respondents were able to address the issues raised in the survey.

Allen & Lambating (2001) previously reported results of descriptive analysis on the responses to the first two survey questions without data from high school students. The focus of this paper is based on the remaining two survey questions and further analysis of questions 1 and 2 in relation to responses to questions 3 and 4 along with new information gathered from high school students.

Research Issue #1

How are the perspectives of high school students, full-time high school teachers, undergraduate and graduate students in pre-service teacher education programs, and education professors similar or different regarding how a high school teacher assigns grades?

To investigate this question frequency distributions and a chi-square analysis were used to analyze the survey question: "Should Sarah change James' grade? Why/why not?" There were 597 valid cases out of 606 for this item. Table 1 summarizes the subjects' responses.

Table 1. Response to the question "Should Sarah change the grade of James?"

| Respondents | N | Yes (%) | No (%) |
|---|---|---|---|
| Undergraduates | 153 | 14 (9%) | 139 (91%) |
| Practicing teachers | 81 | 11 (14%) | 70 (86%) |
| Graduate students | 49 | 7 (14%) | 42 (86%) |
| Professors of education | 34 | 11 (32%) | 23 (68%) |
| High school students | 280 | 23 (8%) | 257 (92%) |
| Overall | 597 | 66 | 531 |

Based on the data, we can see that there is consistency among the five different groups that Sarah should not change the grade. The number of respondents in favor of changing her grade is consistently lower than the number of respondents who responded that Sarah should not change the grade. In looking closely at the high school student group, 92 % of the students responded

"no" to this question compared to an overall 89% of the total number of respondents. A chi-square test was computed to determine if the two variables "group" and "change grade" are independent of each other or not. Results show that there is a significant interaction between the two variables (chi-square = 19.592, $p < .01$). Professors of education are more likely to change the grade of James compared to the other groups.

The data for those respondents that indicated that the grade should not be changed were analyzed by the frequency that a particular reason was given for each group and are summarized in the following tables. As Table 1.1 indicates the high school students had the lowest response rate to citing homework as a requirement as reason for not changing the grade. All other groups except for the professors and the high school students had percentages over 90.

Table 1.1 Grades should not be changed because homework was a requirement.

| | Number of respondents giving this as reason | % |
|---|---|---|
| Undergraduates | 133 | 96 |
| Practicing teachers | 65 | 93 |
| Graduate students | 40 | 95 |
| Professors of education | 21 | 88 |
| High school students | 212 | 82 |
| overall | 471 | 89 |

There is a big difference in the perception of grading between the high school students and the other groups as shown in Table 1.2. Ten percent or less for each of the group of educators cited "changing the grade would be unfair to others" while 32% of the high school students wrote it as one of the reasons for not changing the grade of James.

Table 1.2 Grades should not be changed because it is unfair to others.

| | Number of respondents giving this as reason | % |
|---|---|---|
| Undergraduates | 10 | 7 |
| Practicing teachers | 4 | 6 |
| Graduate students | 4 | 10 |
| Professors of education | 2 | 8 |
| High school students | 82 | 32 |
| overall | 102 | 19 |

Research Issue #2

Do high school students, full-time high school teachers, undergraduate and graduate students in pre-service teacher education programs, and education professors give more or less weight to completing an assigned task (e.g., homework) than demonstrating mastery of the subject in the assignment of grades?

To investigate this question frequency distributions were used to analyze the two survey questions:

- What do you believe should be the purpose of homework?
- What do you believe should be the purpose of a grade?

As indicated in Table 2, the percentage of respondents across the groups had "practice" as the most frequently given purpose for homework. For the practicing teachers, this was followed closely by "reinforcement" and then for "achievement/enrichment". For the undergraduates, "feedback" was next to "practice" then followed by "reinforcement". For the high school students "feedback" was the least frequent reason given.

Table 2. Purpose of homework (in %)

| | Undergraduates (n = 153) | Practicing teachers (n= 81) | Graduate students (n=49) | Professors of education (n= 34) | High school students (n = 280) |
|---|---|---|---|---|---|
| practice | 61 | 68 | 61 | 65 | 58 |
| tracking | 1 | | 6 | | 25 |
| reinforcement | 37 | 51 | 20 | 32 | 23 |
| achievement/ enrichment | 22 | 33 | 16 | 32 | 33 |
| feedback | 46 | 28 | 49 | 29 | 11 |

As Table 2.1 indicates both educators and high students agree that homework is for practice (63% and 58%) and yet this was not reflected in their response to whether the grade of James should be changed or not. One would think that if homework is seen as "practice," then it would be difficult to use it to measure academic achievement in a valid and reliable manner. It should be noted that responses to survey question 1 [Should Sarah change James' grade? Why/why not?] show that a great majority of the respondents (91% Undergraduates, 86% Graduates, 86% Practicing Teachers, 68% College Professors, 92% High School Students) would not change the B grade of James since he was not doing his homework, even though he had a 98% test average (Table 1).

Table 2.1 High school students' response to "purpose of homework" compared to all Educators.

| | Educators | High school |
|---|---|---|
| Practice | 63 % | 58 % |
| Tacking | 2 % | 25 % |
| Reinforcement | 37 % | 23 % |
| Achievement/enrichment | 25% | 33% |
| Feedback | 40% | 11% |
| Discipline and responsibility | 26% | 10 % |
| Test preparation | 6% | 14 % |

Responses to the fourth survey question [What do you believe should be the purpose of a grade?] are summarized in Tables 2.2, 2.3, & 2.4. The data show that 28% of the Educators see the purpose of a grade as a measure of "what was learned", 26% as a combined measure of "knowledge, effort and class involvement", 23% as a measure of "overall performance", 22% as a measure for "mastery", 8% for "placement and feedback" purposes, and 4% to "provide a goal". However, for the High School Students the combination of multiple factors ["knowledge, effort and involvement"] was the reason given for the purpose of a grade (39%). This reason was given twice as often as any other reason. Looking at the terminology used in respondents' descriptions of the purpose of a grade, we can see that there is a similarity between "what was learned" (Table 2.3) and "mastery" (Table 2.4). It is important to note this similarity because it would imply that grades as a measure of what a student knows of the subject matter could possibly be considered a major purpose by more educators than a cursory look at the above numbers show. If these two reasons are combined for Educators the percentages range between 45% and 57%, but total only 25% for High School Students. What is problematic is in the overlap between "what is learned"/"mastery" and "knowledge, effort and involvement". The inclusion of effort and involvement makes it hard for the receiver of the information to interpret the meaning of the grade. Looking further at the data, a discriminant function analysis was done on the question 1 variable 'change grade' [yes/no] as criterion with 'homework was a requirement', 'group', 'age', 'grade level', 'years of teaching', and 'what grade' is deserved (survey question 2) as predictors. Results of the analysis show that whether someone will 'change the grade' is directly correlated to the predictor was 'homework a requirement' and inversely with 'what grade' is deserved.

Table 2.2. What is the purpose of grade?

| | Educators | High school |
|---|---|---|
| To measure what has been learned | 28% | 6% |
| To measure mastery | 22% | 19% |
| Overall performance | 23% | 16% |
| Placement/feedback | 8% | 13% |
| Provides a goal | 4% | 10% |
| Knowledge, effort and involvement | 26% | 39% |

Table 2.3. Grade as a measure of what a student has learned

| | | | |
|---|---|---|---|
| Undergraduates | 153 | 46 | 30% |
| Practicing teachers | 81 | 16 | 20% |
| Graduate students | 49 | 17 | 35% |
| Professors of education | 34 | 9 | 26% |
| High school students | 280 | 18 | 6% |

Table 2.4. Grade as a measure of mastery

| | | | |
|---|---|---|---|
| Undergraduates | 153 | 32 | 21% |
| Practicing teachers | 81 | 20 | 25% |
| Graduate students | 49 | 11 | 22% |
| Professors of education | 34 | 8 | 24% |
| High school students | 280 | 54 | 19% |

Research Issue #3

Does taking a stand-alone measurement course make a difference among educators' perspectives on grading?

The groups using "having taken a measurement course" as variable was cross tabulated with the response to the first survey question, "Should Sarah change her grade?" These groups included the professors of education, practicing classroom teachers, graduate, and undergraduate college students. Table 3 shows that there are a significant number of educators (63%), including many practicing teachers (46%), who have not taken a measurement course.

Table 3. Educators from each group with a measurement course.

| Group | Taken a course | Not taken a course | Total |
|---|---|---|---|
| Undergraduates | 22 | 130 | 152 |
| Practicing teachers | 43 | 37 | 80 |
| Graduate students | 21 | 28 | 49 |
| Professors of education | 30 | 4 | 34 |
| Total | 116 | 199 | 315 |

Of the 116 educators who had taken a measurement course, 18 (16%) of them responded "yes" and 98 (84%) responded "no" to the question regarding the teacher changing the student's grade. Out of the 199 educators who had not taken a measurement course, 24 (12%) would change the grade and 175 (88%) would not. A chi-square test was calculated to see if the two events, having taken or not a measurement course and changing the grade or not are independent of each other. No significant interaction was found (chi-square (1)=.758, $p > .05$).

Results in Table 3.1 between the two groups indicate that the only noticeable difference is in "HW was a requirement". In all the rest of the reasons cited the percentage in the two groups are almost equal. A chi-square test was also calculated to determine if there is a significant interaction between having taken a measurement course and citing "requirement" as a reason for not changing the grade. The two variables were found to be significantly dependent upon each other (chi-square (1) = 5.286, $p < .05$). Those who have not taken a measurement course were more likely to say that they will not change the grade because homework was a requirement than those who had taken a measurement course.

Table 3.1 Reasons given as to why grade should not be changed.

| James grade should not be changed because ... | Taken a course (n=116) in percent | Not taken a course (n=199) in percent |
|---|---|---|
| HW was a requirement | 75 | 85 |
| Unfair to others | 8 | 6 |
| Will give a wrong impression | 6 | 5 |
| Will mean changing grade of all students | 3 | 2 |
| Did not meet requirements | 52 | 51 |
| Must suffer consequence | 7 | 7 |

Results in Table 3.2 between the two groups indicate that there was a higher percentage of educators who had taken a measurement course that placed more importance on "knowing the material' and "contribution to class" than those who did not have a measurement course. On the other hand, more educators who had not taken a measurement course put more value on "engagement of student" in determining the grade.

Table 3.2 Reasons given as to why the change should be changed

| James grade should be changed because... | Taken a course (n=116) in percent | Not taken a course (n=199) in percent |
|---|---|---|
| Unclear policy | 7 | 5 |
| Contributes to class | 10 | 6 |
| Knows material | 16 | 13 |
| Student was engaged | 1 | 3 |

Table 3.3 illustrates areas where there appears to be a difference between the two groups regarding the purpose of homework. Differences were found for "Practice" (6%), "Reinforcement" (9%), "Feedback" (19%), and "Achievement/responsibility" (13 %). A significant interaction was found between "having taken a measurement course" and "practice" as purpose of homework (chi-square = 9.22, $p < .05$). Those who have taken measurement

course were less likely to say that homework is for practice than those who had not taken a measurement course.

Table 3.3 Purpose of homework

| Homework is for … | Taken a course (n=116) in percent | Not taken a course (n=199) in percent |
|---|---|---|
| Tracking | 1 | 2 |
| Practice | 59 | 65 |
| Reinforcement | 43 | 34 |
| Discipline and responsibility | 23 | 24 |
| Test preparation | 8 | 8 |
| Feedback | 28 | 47 |
| Achievement/enrichment | 34 | 21 |

Table 3.4 indicates that there is more of a tendency for those who have not taken a measurement course to list more than one reason for grading than those who have taken a measurement course. There is a greater overlap in the values in column 3 than in column 2. Column 2 sums to 103 but column 3 sums to 116. It is worth noting however that there is almost an equal percentage of educators from the two groups who listed "mastery" as a reason for grading. Significant interaction was also found between "having taken a measurement course" and indicating that the purpose of grading was to measure "mastery"(chi-square (1)= 17.538, $p < .05$). Significance was also found between the variables "having taken a measurement course" and indicating that the purpose of grading was to measure "overall performance" (chi-square (1) = 22.953, $p < .05$). In both cases those educators who had not taken a measurement course were more likely to name those reasons than educators who had taken a measurement course.

Table 3.4 What is the purpose of grading?

| The grade is for ... | Taken a course (n=116) in percent | Not taken a course (n=199) in percent |
|---|---|---|
| What was learned | 25 | 29 |
| Knowledge, effort and involvement | 24 | 28 |
| Mastery | 22 | 23 |
| Overall performance | 19 | 26 |
| Placement | 10 | 6 |
| Provides a goal | 3 | 4 |

Educational Importance

The importance of providing teachers with the necessary background knowledge on classroom assessment, grading procedures and the complexity of factors surrounding the issue of grades can not be underestimated. It is particularly important given the universal use of grades as a commodity for advancement and opportunities throughout our education system. As an academic assessment measure, a grade should have the characteristics of any good measure, namely reliability and validity. Validity refers to the soundness, or accuracy, of the assignment, as well as the interpretation made of the grade. This interpretation should correspond and be dependent on appropriate criteria used by teachers in giving grades to students. Therefore, it is very important for teachers to see that the final (end of the semester) grades primarily reflect what the student knows of the subject at hand. The teacher needs to have enough knowledge on the nature of students and of students' learning to be able to successfully use the other non-achievement functions grades often serve (e.g., motivation, behavioral change, feedback) within the context of formative assessment rather than invalidating the final grade as a summative measure of student achievement.

References

Allen, J.D. (1999). Teaching with cases and lessons learned: The journey of an educational psychologist. In M. R. Sudzina (Ed.) *Case applications for teacher education: Cases of teaching and learning in the content areas*, pp. 36-55. Needham Heights, MA: Allyn & Bacon.

Allen, J. D. (1995). A research study to investigate the development of reflective thought processes of preservice teachers through the use of case studies in educational psychology courses. In H.E. Klein (Ed.) *Case method research & applications: Vol. VII. The art of interactive teaching* (pp. 3-14). Needham, MA: WACRA.

Allen, J. D. & Lambating, J. (2001). *Validity and reliability in assessment and grading: Perspectives of preservice and inservice teachers and teacher education professors.* Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, April, 2001.

*American Heritage Dictionary* (1983). New York: Dell

Barnes, L. B., Bull, K. S., Campbell, N. J., & Perry, K. M. (1998). *Discipline-related differences in teaching and grading philosophies among undergraduate teaching faculty.* Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, April, 1998.

Barnes, S. (1985). A study of classroom pupil evaluation: The missing link in teacher education. *Journal of Teacher Education, 36* (4), 46-49.

Baron, P. A. B. (2000). *Consequential validity for high school grades: What is the meaning of grades for senders and receivers?* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, April, 2000.

Brookhart, S. M. (1998). *Teaching about grading and communicating assessment results*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Diego, April, 1998.

Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7* (4), 279-301.

Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement, 30* (2), 123-142.

Cross, L. H. & Frary, R. B. (1996). *Hodgepodge grading: Endorsed by students and teachers alike*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New York, April, 1996.

Friedman, S. J. & Frisbie, D. A. (1995). The influence of report cards on the validity of grades reported to parents. *Educational and Psychological Measurement, 55*(1), 5-26.

Gage, N. L. & Berliner, D. C. (1991). *Educational psychology, 5th ed.* Boston: Houghton Mifflin Co.

Gall, J. P., Gall, M. D. & Borg, W. (1999). *Applying educational research: A practical guide, 4th ed.*, New York: Addison Wesley Longman.

Gallagher, J. D. (1998). *Classroom assessment for teachers*. Upper Saddle River, NJ: Merrill/Prentice Hall.

Goodwin, A. L. (2001). *The case of one child: Making the shift from personal knowledge to informed practice.* Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, April, 2001.

Gredler, M.E. (1999). *Classroom assessment and learning*. New York: Longman.

Gronlund, N. E. (1998). *Assessment of student achievement*. Needham Heights, MA: Allyn & Bacon.

Jasmine, T. (1999, May). Grade distributions, grading procedures, and students' evaluations of instructors: A justice perspective. *Journal of Psychology, 133* (3), 263-271.

Johnsen, S. (1995). How did you do? The myths and facts about grades. *Gifted Child Today Magazine, 18* (2), 10-13.

Kain, D. L. (1996). Looking beneath the surface: Teacher collaboration through the lens of grading practices. *Teachers College Record, 97*(4), 569-587.

Linn, R. L. & Gronlund, N.E. (1995). *Measurement and assessment in teaching, 7th ed.*, Englewood Cliffs, NJ: Merrill/Prentice Hall.

Lomax, R. G. (1996). On becoming assessment literate: An initial look at preservice teachers' beliefs and practices. *Teacher Educator, 31* (4), 292-303.

Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.

Manzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.

McMillan, J. H. & Schumacher, S. (2001). *Research in education: A conceptual approach, 5th ed.*, New York: Addison Wesley Longman.

Mertens, D. M. (1998). *Research methods in education and psychology: Integrating diversity with quantitative & qualitative approaches*. Thousand Oaks, CA: Sage.

Olson, G. H. (1989). *On the validity of performance grades: The relationship between teacher-assigned grades and standard measures of subject matter acquisition*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, March, 1989.

Oosterhof, A. (2001). *Classroom application of educational measurement*. Upper Saddle River, NJ: Prentice-Hall.

Pintrich, P. R. & Schunk, D. H. (2002). *Motivation in Education, 2nd ed.* Upper Saddle River, NJ: Merrill/Prentice Hall.

Popham, W. J. (2000). *Modern educational measurement: Practical guide for educational leaders, 3rd ed.* Needham, MA: Allyn and Bacon.

Popham, W. J. (1999). *Classroom assessment: What teachers need to know.* Boston: Allyn and Bacon.

Richert, A. E. (1991). Case methods and teacher education: Using cases to teach teacher reflection. In B. R. Tabachnick & K. E. Zeichner (Eds.), *Issues and practices in inquiry-oriented teacher education*, (p. 130-150). London: The Falmer Press.

Silverman, R, Welty, W. & Lyon, S. (1996). *Case studies for teacher problem solving, 2nd ed.*, New York: McGraw-Hill.

Snowman, J., Biehler, R. F., & Bonk, C. J. (2000). *Psychology applied to teaching, 9th ed.*, Boston: Houghton Mifflin.

Stiggins, R. J. (2001). *Student-involved classroom assessment, 3rd. ed.* Upper Saddle River, NJ: Merrill/Prentice Hall.

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, 18*(1), 23-27.

Stiggins, R. J. (1988). Make sure your teachers understand student assessment. *Executive Educator, 10* (8), 24-26.

Sudzina, M. R. (Ed.). (1999). *Case study applications for teacher education: Cases of teaching and learning in the content areas.* Needham Heights, MA: Allyn & Bacon.

Tombari, M. & Borich, G. (1999). *Authentic assessment in the classroom.* Upper Saddle River, NJ: Merrill/Prentice Hall.

# Appendix A

## Percentage of undergraduate and graduate education programs requiring/not requiring a measurement or assessment course at a random sample (n) of colleges and universities

| | N* | n | # of programs | No | Yes | Yes | Yes |
|---|---|---|---|---|---|---|---|
| US Colleges | 1489 | 16 | 47 | 53% | 32% | 4% | 11% |
| CA Colleges | 119 | 13 | 26 | 73% | 15% | 7% | 4% |
| NY Colleges | 135 | 15 | 45 | 67% | 20% | 0% | 13% |

No = No measurement/assessment course required
Yes = General Measurement/assessment course required
Yes = Measurement/assessment course required - focus on informal assessment
Yes = Measurement/assessment course required - focus on students with special needs

*SOURCE: http://www.utexas.edu/world/univ/state/
Web U.S. Universities, by State
"This page of UT Austin Web Central contains a list of regionally-accredited U.S. universities organized by state. ... This list includes only 4-year institutions."

## Appendix B

### Directions

This is an anonymous survey. DO NOT place your name on it. Please keep all of the materials attached. On this first section of the survey please complete all the following items that are relevant to you and your background. **Circle** all the information that best describes you.

- **Age:** 11-14 15-18 19-22 22-26 27-30 31-34 35-39 40-43 44-47 48-51 52+
- **Sex:** Female Male
- **Current grade level (K-12):** K-6th student Middle/Junior H. S. student H. S. student
  K-6th teacher Middle/Junior H.S. teacher H.S. teacher
  Reading teacher Special Ed teacher Administrator
- **Current school level (college):** Undergraduate Student Graduate Student Faculty
- **I have been a K-12 teacher for:** 1-3 years 4-6 yr. 7-9 yr. 10+ yr.
- **I have been a K-12 administrator for:** 1-3 years 4-6 yr. 7-9 yr. 10+ yr.
- **I have been a college teacher for:** 1-3 years 4-6 yr. 7-9 yr. 10+ yr.
- **College Major/Department to which you are associated:**

  Music Education Art Education Early Childhood Ed
  Elementary Education Special Ed/Elem. Ed (SEED) Special Education
  English/Secondary Ed Math/Secondary Ed Biology/Secondary Ed
  Social Studies/Secondary Ed Spanish/Secondary Ed Chemistry/Secondary Ed
  Educational Psychology School Psychology Communication Disorders (CMD)
  Ed Administration Reading Counseling Applied Technology Ed (ATE)

- **I have taken a Graduate level course on educational tests and measurement:** No Yes
  - **I have taken EPY529: Mental & Educational Measurement at St. Rose:** No Yes
- **I have taken an Undergraduate level course on educational tests and measurement:** No Yes

### Directions

Please carefully read the attached case study, *Sarah Hanover*[2]. After reading it, answer the questions on the following pages as concisely as possible. If you find it necessary, you may attach additional comments.

---

[2] Silverman, R., Welty, W. & Lyon, S. (1996). *Case studies for teacher problem solving, 2nd ed.* New York: McGraw-Hill.

Please answer the following questions based on the assumption that Sarah must record one of the following letter grades for James on his report card/school record [ A, B, C, D, F ].

- Should Sarah change James' grade? No Yes
- Why/Why not? _____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

- What letter grade does James deserve? A B C D F
- Why? _____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

- What do <u>you</u> believe should be the purpose of homework?

- What do <u>you</u> believe should be the purpose of a grade?

Thank you for your time and thoughtful completion of this survey.

Dr. James D. Allen

## Appendix C

1. Should Sarah change James' grade? Yes/No, Why/Why not?

Yes

- classroom policy was unclear
- contribution to class was sufficient
- peer tutoring exhibits effort
- after make-up work/extra assignments
- needs to meet individual needs of James

No

- homework was a requirement
- changing it would be unfair to others
- it would give James the wrong impression
- she would have to change all grades

2. What letter grade does James deserve? Why?

A

- classroom policy unclear
- his engagement in class
- he knows the material

A-/B+

- compromise
- his test average is an A

B

- minus 10% for no homework
- he did not meet the requirements
- he did not meet the requirements
- he must suffer the consequences

3. What do you believe should be the purpose of homework?

- practice
- tracking student progress
- reinforcement
- achievement improves
- feedback
- discipline and responsibility
- preparation for tests and new material
- retention of information
- enrichment and challenge

4. What do you believe should be the purpose of a grade?

- meet requirements
- determine what was learned
- knowledge, effort, involvement
- mastery
- performance overall
- provides a goal
- future placement
- feedback

**U.S. Department of Education**
*Office of Educational Research and Improvement (OERI)*
*National Library of Education (NLE)*
*Educational Resources Information Center (ERIC)*

ERIC®

# REPRODUCTION RELEASE

TM033795

(Specific Document)

## I. DOCUMENT IDENTIFICATION:

| | |
|---|---|
| Title: How the multiple functions of grades influence their validity and value as measures of academic achievement | |
| Author(s): Allen, J.D. & Lambating, J. | |
| Corporate Source: The College of St. Rose & The California State University, Sacramento | Publication Date: April 2002 |

## II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1 ↑ ☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A ↑ ☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B ↑ ☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

*I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.*

**Sign here, → please**

| | |
|---|---|
| Signature: James D. Allen | Printed Name/Position/Title: JAMES Allen, Professor of Ed. Psych. |
| Organization/Address: The College of St. ROSE Albany, NY 12203 | Telephone: (518) 454-5273 FAX: (518)-454-2083 |
| | E-Mail Address: allenj@mail.strose.edu Date: 8 April 2002 |

(over)

## III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

| |
|---|
| Publisher/Distributor: |
| Address: |
| Price: |

## IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

| |
|---|
| Name: |
| Address: |

## V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

**ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION**
**UNIVERSITY OF MARYLAND**
**1129 SHRIVER LAB**
**COLLEGE PARK, MD 20742-5701**
**ATTN: ACQUISITIONS**

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

**ERIC Processing and Reference Facility**
**4483-A Forbes Boulevard**
**Lanham, Maryland 20706**

**Telephone: 301-552-4200**
**Toll Free: 800-799-3742**
**FAX: 301-552-4700**
**e-mail: ericfac@inet.ed.gov**
**WWW: http://ericfac.piccard.csc.com**

EFF-088 (Rev. 2/2000)